

## Some Ratio Type Estimators under Measurement Errors

<sup>1</sup>Mukesh Kumar, <sup>1</sup>Rajesh Singh, <sup>2</sup>Ashish K. Singh and <sup>3</sup>Florentin Smarandache

<sup>1</sup>Department of Statistics, Banaras Hindu University (U.P.), India

<sup>2</sup>College of Management Studies, Raj Kumar Goel Institute of Technology, India

<sup>3</sup>Department of Mathematics, University of New Mexico, Gallup, USA

**Abstract:** This article addresses the problem of estimating the population mean using auxiliary information in the presence of measurement errors. A comparative study is made among the proposed estimators, the mean per unit estimator and the ratio estimator in the presence of measurement errors.

**Key words:** Auxiliary variate • Measurement error • Observational error • Mean square error • Efficiency

### INTRODUCTION

In survey sampling, the properties of the estimators based on data usually presuppose that the observations are the correct measurements on characteristics being studied. Unfortunately, this ideal is not met in practice for a variety of reasons, such as non response errors, reporting errors and computing errors. When the measurement errors are negligible small, the statistical inferences based on observed data continue to remain valid. On the contrary when they are not appreciably small and negligible, the inferences may not be simply invalid and inaccurate but may often lead to unexpected, undesirable and unfortunate consequences [1]. Some authors including [2-7] have paid their attention towards the estimation of population mean  $\mu_y$  of the study variable  $y$  using auxiliary information in the presence of measurement errors.

For a simple random sampling scheme, let  $(x_i, y_i)$  be observed values instead of the true values  $(X_i, Y_i)$  on two characteristics  $(x, y)$  respectively for the  $i^{\text{th}}$  ( $i=1, 2, \dots, n$ ) unit in the sample of size  $n$ . Let the measurement errors be

$$u_i = y_i - Y_i \quad (1.1)$$

$$v_i = x_i - X_i \quad (1.2)$$

which are stochastic in nature with mean zero and variances  $\sigma_u^2$  and  $\sigma_v^2$  respectively and are independent.

Further, let the population means of  $(x, y)$  be  $(\mu_x, \mu_y)$  population variances of  $(x, y)$  be  $(\sigma_x^2, \sigma_y^2)$  and  $\sigma_{xy}$  and  $\rho$

be the population covariance and the population correlation coefficient between  $x$  and  $y$  respectively [4].

In this paper we study the behaviour of some estimators in presence of measurement error.

**Exponential Ratio-type Estimator under Measurement Error:** [8], suggested an exponential ratio type estimator for estimating  $\bar{Y}$  as

$$t_1 = \bar{y} \exp \left( \frac{\mu_x - \bar{x}}{\mu_x + \bar{x}} \right) \quad (2.1)$$

$$\text{Let } w_u = \frac{1}{\sqrt{n}} \sum_{i=1}^n u_i, w_y = \frac{1}{\sqrt{n}} \sum_{i=1}^n (y_i - \mu_y)$$

$$w_v = \frac{1}{\sqrt{n}} \sum_{i=1}^n v_i, w_x = \frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \mu_x)$$

$$C_x = \frac{\sigma_x}{\mu_x} \text{ and } C_y = \frac{\sigma_y}{\mu_y}$$

Expression (2.1) can be written as

$$t_1 = \left[ \mu_y + (\bar{y} - \mu_y) \right] \exp \left[ \frac{\mu_x - (\mu_x + \bar{x} - \mu_x)}{\mu_x + (\mu_x + \bar{x} - \mu_x)} \right] \\ (\mu_y + k_1) \exp \left[ \frac{-k_2}{2\mu_x + k_2} \right] \quad (2.2)$$

where,

$$k_1 = \bar{y} - \mu_y = \frac{1}{\sqrt{n}} (w_y - w_u) \text{ and } k_2 = \bar{x} - \mu_x = \frac{1}{\sqrt{n}} (w_x - w_v) \quad (2.3)$$

On simplifying expression (2.2), we have

$$\begin{aligned} (t_1 - \mu_y) &= \mu_y \left[ -\frac{1}{2} \left( \frac{k_2}{\mu_x} \right) + \frac{3}{8} \left( \frac{k_2}{\mu_x} \right)^2 + \dots \right] \\ &+ k_1 \left[ 1 - \frac{1}{2} \left( \frac{k_2}{\mu_x} \right) + \frac{3}{8} \left( \frac{k_2}{\mu_x} \right)^2 + \dots \right] \end{aligned} \quad (2.4)$$

On taking the expectations and using the results

$$E(k_1) = E(k_2) = 0$$

$$E(k_1^2) = \frac{\sigma_y^2}{n} \left( 1 + \frac{\sigma_u^2}{\sigma_y^2} \right) = V_{ym} \quad (2.5)$$

$$E(k_2^2) = \frac{\sigma_x^2}{n} \left( 1 + \frac{\sigma_v^2}{\sigma_x^2} \right) = V_{xm} \quad (2.6)$$

$$E(k_1 k_2) = \frac{\rho \sigma_y \sigma_x}{n} = V_{yxm} \quad (2.7)$$

Taking expectation on both side of (2.4) the bias of  $t_1$  to first order of approximation is

$$\text{Bias}(t_1) = \frac{1}{\mu_x} \left( \frac{3}{8} R_m V_{xm} - \frac{1}{2} V_{yxm} \right) \quad (2.8)$$

where  $R_m = \frac{\mu_y}{\mu_x}$ .

Squaring both side of (2.4) and taking expectation, the mean square error of  $t_1$  up to the first order of approximation, is

$$\begin{aligned} \text{MSE} &= E(t_1 - \mu_y)^2 \\ &= \frac{\mu_y^2}{4\mu_x^2} E(k_2^2) + E(k_1^2) - \frac{\mu_y}{\mu_x} E(k_1 k_2) \\ &= \frac{1}{4} R_m^2 \left( \frac{\sigma_x^2}{n} \left( 1 + \frac{\sigma_v^2}{\sigma_x^2} \right) \right) + \frac{\sigma_y^2}{n} \left( 1 + \frac{\sigma_u^2}{\sigma_y^2} \right) - \frac{\rho \sigma_y \sigma_x}{n} R_m \\ &= \frac{\sigma_y^2}{n} \left[ 1 - \frac{C_x}{C_y} \left( \rho - \frac{C_x}{4C_y} \right) \right] + \frac{1}{n} \left[ \frac{\mu_y^2}{4\mu_x^2} \sigma_v^2 + \sigma_u^2 \right] = M_{t_1}^* + M_{t_1} \end{aligned} \quad (2.9)$$

where  $M_{t_1}^* = \frac{\sigma_y^2}{n} \left[ 1 - \frac{C_x}{C_y} \left( \rho - \frac{C_x}{4C_y} \right) \right]$  is the mean squared error of  $t_1$  without measurement error and

$M_{t_1} = \frac{1}{n} \left[ \frac{\mu_y^2}{4\mu_x^2} \sigma_v^2 + \sigma_u^2 \right]$  is the contribution of measurement errors in  $t_1$

**Another Estimator under Measurement Error:** [9], suggested a regression type estimator  $t_2$  as-

$$t_2 = \omega_1 \bar{y} + \omega_2 (\mu_x - \bar{x}) \quad (3.1)$$

where  $\omega_1$  and  $\omega_2$  are constants that have no restriction.

Expression (3.1) can be written as

$$t_2 - \mu_y = \omega_1 k_1 (\omega_1 - 1) - \omega_2 \omega_x \quad (3.2)$$

Taking expectation of both side of (3.2), we get the bias of the estimator  $t_2$  to order  $O(n^{-1})$  as

$$\text{Bias}(t_2) = \mu_y (\omega_1 - 1)$$

Squaring both side of (3.2) and taking expectation, the MSE of  $t_2$  to the order  $O(n^{-1})$  is

$$\begin{aligned} \text{MSE}(t_2) &= \mu_y^2 (\omega_1 - 1)^2 + \frac{1}{n} \omega_1^2 \sigma_y^2 + \frac{1}{n} \omega_2^2 \sigma_x^2 - \\ &\frac{2}{n} \omega_1 \omega_2 \rho \sigma_y \sigma_x + \frac{1}{n} (\omega_1^2 \sigma_u^2 + \omega_2^2 \sigma_v^2) \end{aligned} \quad (3.4)$$

$$= M_{t_2}^* + M_{t_2} \quad (3.5)$$

Where,

$M_{t_2}^* = \mu_y^2 (\omega_1 - 1)^2 + \frac{1}{n} \omega_1^2 \sigma_y^2 + \frac{1}{n} \omega_2^2 \sigma_x^2 - \frac{2}{n} \omega_1 \omega_2 \rho \sigma_y \sigma_x$ , is the MSE of  $t_2$  without measurement error and  $M_{t_2} = \frac{1}{n} (\omega_1^2 \sigma_u^2 + \omega_2^2 \sigma_v^2)$  is the contribution of measurement error in  $t_2$ .

Expressing MSE of  $t_2$  as

$$\text{MSE}(t_2) = (\omega_1 - 1)^2 \mu_y^2 + \omega_1^2 a_1 + \omega_2^2 a_2 + 2\omega_1 \omega_2 (-a_3) \quad (3.6)$$

where,

$$a_1 = (V_{ym}), \quad a_2 = (V_{xm}), \quad \text{and} \quad a_3 = (V_{yxm}),$$

Now, optimising MSE of the estimator  $t_2$  with respect to  $\omega_1$  and  $\omega_2$  we get

where,

$$b_1 = \mu_y^2 + a_1, \quad b_2 = -a_3, \quad b_3 = a_2, \quad \text{and} \quad b_4 = \mu_y^2.$$

Using the values of  $\omega_1^*$  and  $\omega_2^*$  from equation (3.7) into equation (3.6), we get the minimum MSE of the estimator  $t_2$  as

$$MSE(t_2)_{\min} = \left[ \mu_y^2 - \frac{b_3 b_4^2}{b_1 b_3 - b_2^2} \right] \quad (3.8)$$

**A General Class of Estimators:** We propose a general class of estimator  $t_3$  as

$$t_3 = [m_1 \bar{y} + m_2 (\mu_x - \bar{x})] \left( \frac{\mu_x}{\bar{x}} \right)^\alpha \exp \left[ \frac{\mu_x - \bar{x}}{\mu_x + \bar{x}} \right]^\beta \quad (4.1)$$

Expanding equation (4.1) and subtracting  $\mu_y$  from both side, we have

$$(t_3 - \mu_y) = \left\{ (m_1 - 1) \mu_y + m_1 \mu_y \left\{ -B \frac{k_2}{\mu_x} + \frac{k_2^2 A}{\mu_x^2 8} \right\} + m_1 k_1 \left\{ 1 - \frac{B k_2}{\mu_x} + \frac{k_2^2 A}{\mu_x^2 8} \right\} \right\} \quad (4.2)$$

Taking expectation of both side of (4.2), we get the bias of the estimator  $t_3$  to the order  $O(n^{-1})$

$$Bias(t_3) = (m_1 - 1) \mu_y + m_1 \mu_y \left\{ \frac{V_{xym} A}{8 \mu_x^2} \right\} - m_1 \left\{ \frac{B}{\mu_x} V_{yxm} \right\} + m_2 \left\{ \frac{B}{\mu_x} V_{xym} \right\} \quad (4.3)$$

where,

$$A = \frac{1}{8} [4\alpha(\alpha + 1) + \beta(\beta + 2) + 4\alpha\beta], \quad \text{and} \quad B = \left( \alpha + \frac{\beta}{2} \right)$$

Squaring both side of (4.2) and taking expectation, the MSE of  $t_3$  to the order  $O(n^{-1})$  is

$$\begin{aligned} MSE(t_3) &= E(t_3 - \mu_y)^2 \\ &= (m_1 - 1)^2 \mu_y^2 + m_1^2 (V_{ym}) + (m_1 B R_m + m_2)^2 (V_{xm}) - 2m_1 (m_1 B R_m + m_2) (V_{yxm}) \\ &\quad - 2(m_1 - 1) \mu_y \left\{ \frac{m_1 B}{\mu_x} V_{yxm} + \frac{1}{\mu_x} \left( \frac{m_1 R_m A}{8} + m_2 B \right) V_{xm} \right\} \\ MSE(t_3) &= m_1^2 (\mu_y^2 + V_{xm} + B^2 R_m^2 - 2B R_m V_{yxm}) + m_2^2 (V_{xm}) + 2m_1 m_2 (B R_m V_{xm} - V_{yxm}) \\ &\quad - 2m_1 \mu_y^2 + \mu_y^2 - 2m_1^2 B R_m V_{yxm} + 2m_1 R_m B V_{yxm} - \frac{m_1^2 R_m^2 A}{4} V_{yxm} \\ &\quad - 2m_2 R_m B V_{yxm} + \frac{R_m^2 A m_1}{4} V_{xm} + 2m_2 R_m B V_{xm} \end{aligned} \quad (4.4)$$

Writing MSE of the estimator  $t_3$  as

$$MSE(t_3) = (m_1 - 1)^2 \mu_y^2 + m_1^2 (A_1 + 2A_3) + m_2^2 A_2 + 2m_1 m_2 (-A_4 - A_5) - 2m_1 A_3 + 2m_2 A_5 \quad (4.5)$$

where,

$$\begin{aligned} A_1 &= (V_{ym} + B^2 R_m^2 V_{xm} - 2R_m B V_{yxm}), \quad A_2 = (V_{xm}), \\ A_3 &= (A R_m^2 V_{xm} - B R_m V_{yxm}), \quad A_4 = (V_{yxm} - B V_{xm} R_m), \\ A_5 &= (-B R_m V_{xm}). \end{aligned}$$

Now, optimising MSE  $t_3$  with respect to  $m_1$  we get the optimum values of  $m_2$  as-

$$m_1^* = \frac{B_3B_4 - B_2B_5}{B_1B_3 - B_2^2} \text{ and } m_2^* = \frac{B_1B_5 - B_2B_4}{B_1B_3 - B_2^2} \quad (4.6)$$

where,

$$B_1 = \mu_y^2 + A_1 + 2A_3, \quad B_2 = -A_4 - A_5, \quad B_3 = A_2, \quad B_4 = \mu_y^2 + A_3 \quad \text{and} \quad B_5 = -A_5.$$

Minimum MSE of the estimator  $t_3$  is given by

$$MSE(t_3)_{\min} = \left[ \mu_y^2 - \frac{B_1B_5^2 + B_3B_4^2 - 2B_2B_4B_5}{B_1B_3 - B_2^2} \right] \quad (4.7)$$

**Theoretical Efficiency Comparison:** The MSE of the proposed estimator  $t_1$  will be smaller than usual estimator under measurement error case, if the following condition is satisfied by the data set

$$\frac{\sigma_y^2}{n} \left[ 1 - \frac{C_x}{C_y} \left( \rho - \frac{C_x}{4C_y} \right) \right] + \frac{1}{n} \left[ \frac{\mu_y^2}{4\mu_x^2} \sigma_y^2 + \sigma_u^2 \right] \leq \frac{\sigma_y^2}{n} \left( 1 + \frac{\sigma_u^2}{\sigma_y^2} \right)$$

or

$$R_m^2 \frac{V_{xm}}{V_{yxm}} \leq 4 \quad (5.1)$$

As the estimator  $t_3$  defined in (3.1) is the particular member of the generalised estimator  $t_3$  defined in (4.1), if the condition (5.2) is satisfied for different values of  $\alpha, \beta$  the  $m_1, m_2$  estimator  $t_3$  (or  $t_2$ ) will be better than usual estimator under measurement errors.

$$MSE(t_3)_{\min} \leq V(\bar{y}_m)$$

if,

$$\left[ \mu_y^2 - \frac{B_1B_5^2 + B_3B_4^2 - 2B_2B_4B_5}{B_1B_3 - B_2^2} \right] \leq V(\bar{y}_m) \quad (5.2)$$

## Empirical Study

**Data Statistics:** The data used for empirical study has been taken from [10].

Where,  $Y_i$  = True consumption expenditure,

$X_i$  = True income,

$y_i$  = Measured consumption expenditure,

$x_i$  = Measured income.

n	$\mu_y$	$\mu_x$	$\sigma_y^2$	$\sigma_x^2$	$\rho$	$\sigma_u^2$	$\sigma_v^2$
10	127	170	1278	3300	0.964	36.00	36.00

Table 6.1: Showing the MSE of the estimators with and without measurement errors.

Estimators	MSE without measurement error	Contribution of measurement error in MSE	MSE with measurement error
Usual estimator ( $\bar{y}$ )	319.500	9.0005	328.5005
$t_1$	64.8134	10.2558	75.0700
$t_{reg}$	22.5000	12.2401	34.7050
$t_{2min}(\alpha = 0, \beta = 0)$	22.4687	12.1619	34.6306
$t_{3min}(\alpha = 1, \beta = 0)$	22.4678	12.1619	34.6306
$(\alpha = 0, \beta = 1)$	22.1033	12.0680	34.1713
$(\alpha = 1, \beta = 1)$	20.9734	12.4075	33.3809

### CONCLUSION

From Table 6.1, we observe that the MSE of the estimator  $t_3$  (for  $\alpha = 1, \beta = 1$ ) is minimum. From third column of the Table we observe that the usual estimator  $(t_1)$  is least affected by measurement errors. We also observe from that Table that, if due care of observational error is not given, the estimate of the variances that we get gives an under estimate of the true variances.

### REFERENCES

1. Srivastava, A.K. and Shalabh, 2001. Effect of measurement errors on the regression method of estimation in survey sampling. *J. Statist. Res.*, 35(2): 35-44.
2. Allen, J., H.P. Singh and F. Smarandache, 2003. A family of estimators of population mean using multiauxiliary information in presence of measurement errors. *Int. J. Soc. Econ.*, 30(7): 837-849.
3. Manisha and R.K. Singh, 2001. An estimation of population mean in the presence of measurement errors. *J. Ind. Soc. Agri. Statist.* 54(1): 13-18.
4. Manisha and R.K. Singh, 2002. Role of regression estimator involving measurement errors. *Brazilian J. Probability Statistics.* 16: 39-46.
5. Shalabh, 1997. Ratio method of estimation in the presence of measurement errors. *J. Ind. Soc. Agri. Statist.* 50(2): 150-155.
6. Singh, H.P. and N. Karpe, 2008. Ratio-product estimator for population mean in presence of measurement errors. *J. Appl. Statist. Sci.*, 16: 49-64.
7. Singh, H.P. and N. Karpe, 2009. On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors. *Statistica, Department of Statistics, University of Bologna*, 69(1): 27-47.
8. Bahl, S. and R.K. Tuteja, 1991. Ratio and product type exponential estimator. *Infrm. and Optim. Sci.* XII, I: 159-163.
9. Koyuncu, N. and C. Kadilar, 2010. On improvement in estimating population mean in stratified random sampling. *Jour. Appl. Stat.*, 37(9): 999-1013.
10. Gujarati, D.N. and Sangeetha, 2007. *Basic Econometrics*. Tata McGraw - Hill.